考虑信任损益的健康大数据开放利用演化博弈分析*

王丹丹，解世程，程英英

河南科技大学商学院　洛阳　471023

摘要：【目的/意义】相关主体之间的信任已成为决定健康大数据开放利用成效的关键因素。基于健康大数据开放利用过程，考虑信任损益因素，研究信任对各方主体行为的影响机制，为提升我国健康大数据开放利用成效提供前瞻性参考。**【方法/过程】**聚焦于数据平台、使用方、患者在考虑信任损益因素情况下的决策行为，基于开放利用过程分析，建立三方演化博弈模型，并通过 MATLAB 进行仿真分析，最后根据均衡点分析与仿真结果提出对策建议。**【结果/结论】**获取患者信任是使用方透明使用数据的根本动力；提升研究环境的安全性可以从根本上解决信任问题；加强监管与增大使用方的信任损益有助于实现良性的开放利用状态；提升患者对数据平台信任水平的同时，加强对使用方行为的制约更为重要。我国可考虑由国家主导建设健康大数据开放利用平台，加强可信研究环境建设并积极推广应用，推行数据使用登记制度并建立标准体系，有效引导公众参与健康大数据协同治理过程。

关键词：健康大数据；信任；透明度；开放利用；演化博弈

中图分类号：G203

健康大数据是指在人们疾病防治、健康管理等过程中产生的与健康医疗相关的数据[1]。美国自 2009 年率先发起开放数据运动，推出政府数据开放平台，并延伸至健康医疗数据领域，建立健康医疗数据开放平台 HealthData.gov，疾病预防中心、医疗保险与医疗补助服务中心、食药监管理局等开放了医疗保险与补助、卫生费用、疾病控制与公共卫生等结构化数据。英国建设了国家医疗服务（National Health Service，NHS）系统，开放人口统计数据、卫生管理与质量监测类数据。依托 NHS 体系，英国在推进健康医疗大数据的开放利用方面走在了世界前列，于 2013 年启动 care.data 项目，但由于数据的不合理使用，致使公众信任丧失，于 2016 年 7 月即宣告停止。2018 年，英国成立国家健康数据研究所（Health Data Research UK，HDR UK），致力于整合英国全域的健康数据以用于科学研究的目的并实施至今[2]。

围绕健康大数据国内学者从权属配置[3-5]、隐私保护[6]、开放利用障碍与治理[7]、质量评价[8]、价值实现路径[9-10]等方面开展研究，国外学者则是从技术进步、法律完善和伦理引导 3 个方面进行探讨，其中探讨技术改进居多，如访问安全[11-12]、存储安全[13]、防范网络攻击[14]、共享安全[15]、处理安全[16]、传输安全[17]、交易安全[18]、构建可信研究环境[19]等。隐私、透明、问责、公平和同意对健康大数据的重要性被充分认同[20-22]，利益相关者之间的信任[23-25]构建被高度关注。我国高度重视医疗健康大数据的开放利用，《关于促进和规范健康医疗大数据应用发展的指导意见》指出要推动数据共享，建立健全开放制度。《国家健康医疗大数据标准、安全和服务管理办法（试行）》确立了以国家为主导的数据管理与开发利用方针。

---

*本文系国家社会科学基金项目“健康大数据的开放利用与协同治理研究”（编号：23BTQ068）的阶段性研究成果之一。

作者简介：王丹丹（ORCID：0000-0001-9214-4906），河南科技大学商学院副院长，教授，博士生导师，Email:wangdan430806@163.com；解世程（ORCID：0009-0006-4471-5018），博士研究生；程英英（ORCID：0000-0003-2078-8782），讲师，博士。

This work is supported by the National Social Science Fund of China project titled“ Research on the Open Utilization and Collaborative Governance of Healthcare Big Data” (Grant No. 23BTQ068) .

Author(s): Wang Dandan, vice dean of the Business School, Henan University of Science and Technology, professor, PhD, doctoral supervisor,E-mail: wangdan430806@163.com; Xie Shicheng, doctoral candidate; Cheng Yingying, lecturer, PhD.

鉴于此，分析英国健康大数据开放利用经验，构建健康大数据开放利用过程模型，考虑在既定的隐私保护技术、制度、基础设施水平下，数据使用方数据使用的透明度、监管方的监管强度与患者参与行为的相互影响，建立三方演化博弈模型，研究各利益相关方的博弈行为并提出对策建议。

1 研究基础/Research basis

1.1 英国健康大数据开放利用历程回顾

NHS 汇聚了英国全域人口从出生到死亡的人类生命历程数据，为英国健康研究提供了丰富的数据基础。英国政府积极推进健康大数据的开放利用，在 NHS 的带领与指导下，2013 年英国卫生和社会保健信息中心（Health and Social Care Information Centre，HSCIC）启动了健康大数据旗舰平台 care.data，从全国范围的公立医疗机构和全科医生诊所收集最详尽的电子健康数据，旨在更好地认识病患、研发药物和治疗方法、预测公共卫生和疾病的发展趋势、监控药物和治疗的安全状况、比较全国各区域的医疗质量等[26]。该项目允许 NHS 和符合条件的私营公司出于上述公共目的使用数据。

在 care.data 项目中，NHS 指导的 HSCIC 为数据收集方与数据使用监管方，数据有偿向科研机构和部分符合条件的私营公司开放，并限制数据使用于公共利益相关研究项目。而实际上，NHS 将数据匿名后出售给卫生组织、研究型大学、医院信托基金、制药公司及其它私营公司后，虽然进行了匿名化并约定使用范围为公共利益，但出现了大量的患者个人信息泄露，并被保险公司、营销公司等用于商业目的，严重侵害了患者的隐私权和自主权。虽然 NHS 随后加强了监管，但由于数据本身的易复制性等特点带来的负外部性，隐私泄露、数据滥用事件并未就此停止。大量公众选择退出该项目，导致 2016 年 7 月 care.data 项目即宣告终止[27]。

care.data 的实践表明：第一，加密技术无法保障数据绝对安全。匿名后数据并非绝对安全，具有可再识别的可能性，仍需安全保护。第二，在无法保障数据绝对安全的情况下，数据流通过程中的透明度和患者同意是保障公众信任的基本前提。例如："家庭医生向 HSCIC 提供数据、HSCIC 对外开放数据等过程，是否需要获得患者的同意？""数据是否用于高质量科研等公共目的而非商业目的？"只有解决好这些问题，才能赢得公众和医护人员的信任，保障数据的良好供给[28]。2018 年，英国成立 HDR UK，致力于以合乎道德的目的使用英国的健康医疗数据[2]。HDR UK 主要通过可信研究环境（Trusted Research Environments，TRE）、数据使用登记（Date Use Registers，DUR）制度、患者与公众参与（Patient and Public Involvement and Engagement，PPIE）制度来建立公众信心、保障公众信任。

1.2 演化博弈理论与信任损益

健康大数据开放利用是多个利益相关方策略选择的动态化过程，涉及公众与患者、数据保管人（医疗机构、信托机构、存储中心等）、数据使用方（医学研究机构、科研院所、制药公司等）、数据监管方（政府机构、数据平台等）以及技术服务方等利益主体，各利益相关方在有限理性的条件下不断试错、进化、不断选择。有学者使用演化博弈方法研究医疗机构、技术支撑企业和政府三方的博弈过程[29]，从隐私保护视角构建了患者和医疗服务机构两方演化博弈模型以及患者、医疗服务机构和政府的三方演化博弈模型，探究了各方利益相关者的共享意愿[30-31]。

信任是一种关系，在这种关系中一方对另一方的能力和善意有合理的信念，或者对执行某项行动有明确的承诺[32-34]。就健康大数据而言，公众信任表现为：①对国家、政府的信任。公众由于不信任自己国家的医学研究水平[35]或政治制度[36-37]而具有较低的共享意愿。②对科研人员、医生的信任。在卫生系统，科研人员的建议已被证明可增加公众参与研究的可能性[38-40]。③对卫生系统与使用方机构的信任。公众与某个实体共享敏感信息的意愿因其信任程度和熟悉程度而不同[41-43]。④对监管机构的信任。公众关注监管数据的个人或组织的声誉

[44-46]。本文将信任损失（或收益）视为一方感知到另一方的能力和善意降低（或升高）时，付出行动后对另一方产生的预期实际利益损失（或收益）。如数据平台强监管时获得的信任收益，即数据平台的强监管行为引起患者感知到源于数据平台的能力和善意的升高时，患者积极参与共享数据而给数据平台带来预期收益。

2 模型构建/Model construction

2.1 问题描述

健康大数据开放利用过程如图 1 所示。数据收集阶段，患者由于治疗需要将数据共享给医疗机构（医院、社区诊所、个体医生等），医疗机构对患者数据拥有控制权，而患者个人仍保留隐私权[3,47]。为促进开放利用，采用选择退出的患者知情同意机制。数据平台将医疗机构的健康医疗数据进行收集，汇聚整理后形成可供利用的健康医疗大数据集。在开放利用阶段，数据平台对数据进行匿名化处理后有偿或无偿向服务于公共利益的数据使用方提供数据，双方签订数据使用协议，约定数据使用范围等，使用方按照协议约定使用数据。

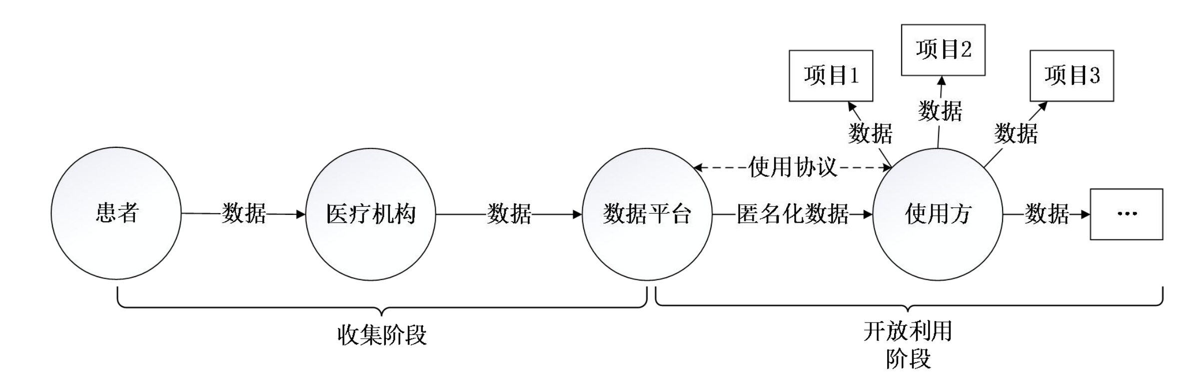


图 1 健康大数据开放利用过程模型

Figure 1 Process model for open utilization of healthcare big data

对于使用方而言，不透明使用数据可能得到更大利益。不透明使用是指使用方除了按照约定的数据使用范围外，对数据再次加工后用于约定以外的项目，或与其它使用者共享[48]。在数据使用申请时虽然明确签订协议限制了使用范围，但由于数据复制成本低，在做好保密工作的情况下使用方数据的再加工或流动不易被察觉。另外，由于收集过程的复杂性，数据已在多处存储，如出现泄露，由于源头的多样性会导致难以快速精确辨认。因此，使用方面临利益诱惑时可能存在侥幸心理不透明使用数据，这会极大增大患者数据被重新识别与泄露的风险。

对于数据平台而言，将数据有偿或无偿提供给使用方可获得数据共享收益或由数据使用方使用数据而带来的社会效益。由于存在数据泄露风险，平台有责任对使用过程进行监管。在隐私保护技术水平既定的情况下，敦促使用方合规使用数据是平台监管的重点，包括数据使用透明度建设与追溯数据泄露源头等。数据使用的透明度建设是一系列工作，需使用方投入一定成本，如敦促实施 PPIE 制度，让公众代表参与数据使用整个过程；敦促实施 DUR 制度，保持数据使用过程公开透明；加强内部的数据访问权限管理与员工培训等。平台对使用方的隐私保护技术水平与透明度建设有最低要求，追查数据泄露事件并采取惩罚措施。

对于患者而言，患者享有不同意、撤销同意以及删除等权利[49-50]，在数据开放利用的整个过程中，患者随时选择是否继续参与。患者参与（或不退出）可能出于获得医疗服务质量提高收益[51]、个人经济利益[52-53]、个人感知的社会公共利益[54-55]等原因。使用方不透明使用数据会增大患者感知的隐私风险，引起患者对数据滥用与隐私泄露的担忧，患者数据隐私敏感性越强，感知的数据泄露损失越大。

2.2 模型假设

数据平台、使用方、患者群体三方博弈主体如图 2 所示。数据平台对监管的强度进行决策，强监管意味着较高的监管成本，同时可获得公众信任收益，此时使用方不透明使用数据时被发现的概率增大，数据平台的策略选择为（强监管，弱监管）；使用方对是否透明使用数据进行决策，透明使用数据意味着严格按照协议使用数据，不透明使用意味着面临不合规使用数据赚取额外收益的诱惑，同时也面临数据泄露或被监管方查处的风险，使用方的策略选择为（透明使用，不透明使用）；患者对是否退出数据共享进行决策，退出意味着不参与，不退出意味着继续参与，患者的策略选择为（参与，不参与）。各参数符号及其含义如表 1 所示。

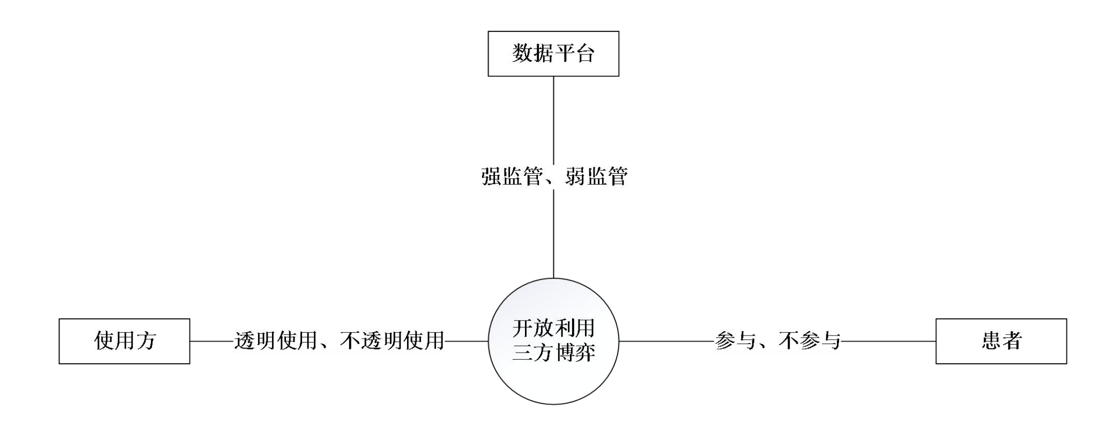


图 2 健康大数据开放利用三方博弈主体

Figure 2   Subjects of the tripartite game of the open utilization

表 1 三方演化博弈的参数符号及其含义

Table 1   Parameters and their meanings for the tripartite evolutionary game

参数符号	参数含义
R_1	患者参与给数据平台带来的直接经济收益与社会效益总和
G	数据平台选择强监管时的成本
T_4	数据平台强监管时的信任收益
k	数据平台强监管时查出使用方不透明使用的概率
F_1	数据平台强监管时对查出的不透明使用方的罚金
F_2	数据泄露事件发生时，数据平台对使用方的罚金
L_1	数据平台弱监管时，数据泄露事件发生时数据平台的信任损失
C	数据使用方的固定投入
R_2	数据使用方透明使用数据情况下的收益
R_3	数据使用方不透明使用数据情况下的额外收益
T_5	数据使用方透明使用数据情况下的信任收益，同时也是不透明使用数据且强监管情况下未被查出也未发生数据泄露事件时的侥幸收益
L_2	数据平台强监管且发生数据泄露事件被追查到时使用方的信任损失
L_3	发生数据泄露事件时弱监管下使用方的信任损失
R_4	患者选择参与时获得的各项收益之和
D	患者参与时发生隐私泄露事件时的损失
a	数据使用方采取透明使用策略时的患者隐私泄露风险
b	数据使用方采取不透明使用策略时的患者隐私泄露风险，$a < b$
T_6	数据平台强监管时，患者的信任收益

假设 1：数据平台、使用方、患者均为有限理性主体。数据平台注重长期可持续发展，考虑信任损益，可视为对未来利益或损失的考虑；使用方在短期利益与长期利益之间进行权衡选择，短期内不透明使用数据可以带来即时利益，但承担信任损失风险，长期内考虑信任损益，可视为对未来利益或损失的考虑；患者对数据平台的信任影响其自身决策，可视为感知的信任损益[56]。

假设 2：以x表示数据平台选择强监管的比例，则 $1-x$为其选择弱监管的比例，$x \in [0,1]$；以y表示使用方选择透明的比例，则 $1-y$为其选择不透明的比例，$y \in [0,1]$；以z表示患者选择参与的比例，则 $1-z$为其选择不参与的比例。

假设 3：数据平台的收益与成本。患者参与时给数据平台带来直接经济收益与社会效益的总和R_1；数据平台强监管时，承担监管成本G，公众对数据开放利用的安全保障更有信心，数据平台获得信任收益T_4[56]；强监管时，使用方不透明使用数据不一定能被查出，设被查出的概率为k，查出后数据平台向使用方收取罚金F_1，且强制使用方中止不合规行为；数据平台强监管时，若数据泄露事件发生，则进行严查，对数据泄露者收取罚金F_2；数据平台强、弱监管为一个相对概念，为简化模型，假设数据平台弱监管时的监管成本忽略不计；数据平台弱监管时，出现数据泄露事件，数据平台信任损失L_1。

假设 4：使用方的收益与成本。使用方必须有既定的隐私保护技术水平与一定程度的透明度建设才有资格使用数据，其固定投入为C；使用方透明使用数据情况下可获得收益R_2，不透明使用数据情况下可获得额外收益R_3；采取透明使用策略时，信任收益为T_5；使用方采取不透明使用策略时，且数据平台强监管时，使用方被查出的概率为k，罚金F_1，并停止不合规行为，发生数据泄漏时使用方被追查到，则罚金F_2，信任损失L_2；采取不透明使用策略且数据平台采取强监管，但未被查出而且未发生数据泄露事件，数据使用方亦可获得侥幸收益T_5；数据平台弱监管时，发生数据泄露事件时无法确认是否使用方承担泄露责任，但会造成信任损失L_3。

假设 5：患者的收益与成本。患者参与时获得的直接收益或感知社会效益等的总和记为R_4；隐私泄露时患者的损失为D；使用方选择透明使用策略时患者隐私泄露风险为a；不透明使用时患者隐私泄露风险为b；由于透明使用数据可降低隐私风险，因此$a<b$；数据平台强监管时，患者感知安全性提升，给患者带来信任收益T_6；患者不参与的净收益为 0，同时也不承担风险。

2.3 收益矩阵

构建数据平台、使用方、患者之间的三方演化博弈收益矩阵，如表 2 所示。

表 2 开放利用阶段三方演化博弈收益矩阵

Table 2 Payoff matrix of the tripartite evolutionary game

数据平台/使用方/患者			数据平台	
		使用方	强监管x	弱监管 $1-x$
患者	参与z	透明使用y	R_1-G+T_4， R_2-C+T_5， R_4-aD+T_6	R_1-aL_1， R_2-C+T_5， R_4-aD
		不透明使用$1-y$	$R_1-G+T_4+kF_1+bF_2(1-k)$， $R_2-C-kF_1-b(1-k)(F_2+L_2)+(1-k)(1-b)(R_3+T_5)$， R_4-bD+T_6	R_1-bL_1， $R_2+R_3-C-bL_3$， R_4-bD
	不参与$1-z$	透明使用y	$-G+T_4$， $-C+T_5$， 0	0， $-C+T_5$， 0
		不透明使用$1-y$	$-G+T_4$， $-C$， 0	0， $-C$， 0

3 模型分析/Model analysis

3.1 复制动态方程

根据表 2，可得到数据平台选择“强监管”和“弱监管”策略的期望收益为U_{11}和U_{12}，平均期望收益U_1分别为：

$$U_{11}=zy(R_1-G+T_4)+z(1-y)[R_1-G+T_4+kF_1+bF_2(1-k)]+y(1-z)(-G+T_4)+(1-z)(1-y)(-G+T_4) \tag{1}$$

$$U_{12}=z(1-y)(R_1-bL_1)+yz(R_1-aL_1) \tag{2}$$

$$U_1=xU_{11}+(1-x)U_{12}=x\{y(z-1)(G-T_4)-z(y-1)[R_1-G+T_4+kF_1-bF_2(k-1)]-(y-1)(z-1)(G-T_4)+yz(R_1-G+T_4)\}-(x-1)[yz(R_1-aL_1)-z(y-1)(R_1-bL_1)] \tag{3}$$

根据动态方程构建法，构建数据平台的复制动态方程为：

$$f(x)=x(U_{11}-U_1)=-x(x-1)(T_4-G+zbF_2+zbL_1+zkF_1-yzbL_1-yzkF_1-zbkF_2-yzbF_2+yzaL_1+yzbkF_2) \tag{4}$$

同理，构建使用方与患者的复制动态方程为：

$$f(y)=y(U_{21}-U_3)=-y(y-1)(T_5-zR_3+zbL_3-xzT_5+xzkF_1+xzbR_3+xzbT_5+xzkR_3+xzkT_5+xzbF_2+xzbL_2-xzbL_3-xzbkF_2-xzbkL_2-xzbkR_3-xzbkT_5) \tag{5}$$

$$f(z)=z(U_{31}-U_3)=-z(z-1)(R_4-bD+xT_6-yaD+ybD) \tag{6}$$

3.2 均衡点稳定性分析

进一步，令$f(x)=0$，$f(y)=0$，$f(z)=0$ 可得到三方博弈的 14 个局部均衡点。由于多群体演化博弈复制动态系统的渐进稳定解一定是严格的纳什均衡[57]，因此只需考虑 8 个纯策略均衡点：$E_1(0,0,0)$，$E_2(0,0,1)$，$E_3(0,1,0)$，$E_4(0,1,1)$，$E_5(1,0,0)$，$E_6(1,0,1)$，$E_7(1,1,0)$，$E_8(1,1,1)$。根据 Friedman 的方法[58]，对$f(x)$、$f(y)$、$f(z)$分别关于x、y、z求偏导，得到复制动态系统的雅可比矩阵 J：

$$
J=\begin{bmatrix}
\begin{gathered}(1-2x)(T_4-G+zbF_2+zbL_1+\\ zkF_1-yzbL_1-yzkF_1-zbkF_2+\\ -yzbF_2+yzaL_1+yzbkF_2)\end{gathered} & \begin{gathered}xz(x-1)(bF_2-aL_1+\\ bL_1+kF_1-bkF_2)\end{gathered} & \begin{gathered}-x(x-1)(bF_2+bL_1+kF_1-\\ bkF_2-ybF_2+yaL_1-ybL_1-\\ ykF_1+ybkF_2)\end{gathered} \\
\begin{gathered}-yz(y-1)(bF_2-T_5+bL_2-\\ bL_3+kF_1+bR_3+bT_5-bkF_2\\ -bkL_2-bkR_3-bkT_5)\end{gathered} & \begin{gathered}(1-2y)(T_5-zR_3+zbL_3-\\ xzT_5+xzkF_1+xzbR_3+xzbT_5\\ +xzkR_3+xzkT_5+xzbF_2+\\ xzbL_2-xzbL_3-xzbkF_2-\\ xzbkL_2-xzbkR_3-xzbkT_5)\end{gathered} & \begin{gathered}-y(y-1)(bL_3-R_3-xT_5\\ +xbF_2+xbL_2-xbL_3+xkF_1\\ +xbR_3+xbT_5+xkR_3+xkT_5\\ -xbkF_2-xbkL_2-xbkR_3-xbkT_5)\end{gathered} \\
-zT_6(z-1) & zD(z-1)(a-b) & \begin{gathered}(1-2z)(R_4-bD+xT_6\\ -yaD+ybD)\end{gathered}
\end{bmatrix}
$$

将 E_1~E_8分别代入雅可比矩阵，计算矩阵特征值，结果如表 3 所示。

表 3 三方演化博弈均衡点稳定性情况

Table 3 Stability of equilibrium points in the tripartite evolutionary game

均衡点	特征值λ_1	特征值λ_2	特征值λ_3	稳定性结论
$E_1(0,0,0)$	T_4-G	T_5	R_4-bD	不稳定点
$E_2(0,0,1)$	$T_4-G+bF_2+bL_1+kF_1-bkF_2$	$T_5-R_3+bL_3$	$bD-R_4$	当各特征值均小于 0 时为 ESS
$E_3(0,1,0)$	T_4-G	$-T_5$	R_4-aD	
$E_4(0,1,1)$	T_4-G+aL_1	$R_3-T_5-bL_3$	$aD-R_4$	
$E_5(1,0,0)$	$G-T_4$	T_5	R_4+T_6-bD	不稳定点
$E_6(1,0,1)$	$G-T_4-bF_2-bL_1-kF_1+bkF_2$	$bF_2-R_3+bL_2+kF_1+bR_3+bT_5+kR_3+kT_5-bkF_2-bkL_2-bkR_3-bkT_5$	$bD-T_6-R_4$	当各特征值均小于 0 时为 ESS
$E_7(1,1,0)$	$G-T_4$	$-T_5$	R_4+T_6-aD	
$E_8(1,1,1)$	$G-T_4-aL_1$	$R_3-bF_2-bL_2-kF_1-bR_3-bT_5-kR_3-kT_5+bkF_2+bkL_2+bkR_3+bkT_5$	$aD-T_6-R_4$	

根据李雅普诺夫第一法[59]可知，只有当雅各比矩阵的所有特征值均为负实部时，得到的均衡点为渐进稳定点（ESS）；当雅各比矩阵的特征值有一个及以上有正实部，得到的均衡点为不稳定点；当雅各比矩阵的特征值有实部为 0 和负实部两种情形时，得到的均衡点处于临界状态，其稳定性不能根据特征值符号确定。据此得出如下结论：

结论 1：对于$E_1(0,0,0)$和$E_5(1,0,0)$，由于$\lambda_2=T_5\geq 0$，因此为不稳定点。T_5为使用方对信任收益的考虑，如果使用方不考虑信任收益或者出于短期利益造成信任损失，即假如T_5＜0，则E_1和E_5将为局部均衡点，会出现无论监管方监管强度如何，使用方会选择不透明使用数据而同时患者会大量退出。此不稳定点表明，正是由于使用方对信任收益的考虑，使得系统避免了（监管/不监管，不透明，不参与）的不利于开放利用的局部均衡状态。反之，如果数据滥用成风，社会信任缺失，那么出现“不想、不敢”共享数据也就不足为奇。

结论 2：对于$E_3(0,1,0)$和$E_7(1,1,0)$，当满足$\lambda_1=T_4-G$＜0 且$\lambda_2=-T_5$＜0 且$\lambda_3=R_4-aD$＜0 时，$E_3(0,1,0)$为 ESS，当满足$\lambda_1=G-T_4$＜0 且$\lambda_2=-T_5$＜0且$\lambda_3=R_4+T_6-aD$＜0 时，$E_7(1,1,0)$为 ESS。但由于在实际情境中，当患者选择不参与策略时，使用方与数据平台的策略选择实际上已经失去了意义，对于患者来说其个人继续参与的风险过大，潜在损失大于其个人收益，即使使用方选择透明使用策略，患者也会选择不参与策略。反映出在既定的

隐私保护技术水平与法律保障下，当患者感知的隐私保护、个人基本权利受损风险过大时，使用方的透明度策略与监管方的监管策略对患者的行为几乎不会产生影响。一个基本事实是现实情境中，数据敏感性极强的患者共享数据的意愿往往极低，例如 care.data 案例中指出的艾滋病毒状况、性传播感染状况、堕胎记录、试管婴儿记录、婚姻状况、投诉记录、犯罪和虐待记录等患者信息不应被收集[26]。

结论 3：对于$E_2(0,0,1)$，当满足$\lambda_1 = T_4 - G + bF_2 + bL_1 + kF_1 - bkF_2$＜0 且$\lambda_2 = T_5 - R_3 + bL_3$＜0 且$\lambda_3 = bD - R_4$＜0 时，$E_2(0,0,1)$为 ESS。这种情形也属于一种极端情形，即在既定的技术水平与法律保障下，当患者面临的隐私保护、个人基本权利受损风险很小时，使用方的透明度策略与监管方的监管策略对患者的行为也几乎不会产生影响。可以设想一种理想情境：如果具备了绝对可信的研究环境，隐私保护技术可以确保隐私安全，法律规制颇为完善，那么数据的开放利用过程中就不会出现隐私泄露等侵权问题。此时，由于患者继续参与倾向显著，那么对于数据平台来说强监管的信任收益T_4较低，对于使用方来说透明使用的信任收益T_5较低，也就不存在监管与透明使用的必要了，数据平台会选择弱监管策略，使用方会选择不透明使用策略。

结论 4：对于$E_4(0,1,1)$，当满足$\lambda_1 = T_4 - G + aL_1$＜0 且$\lambda_2 = R_3 - T_5 - bL_3$＜0 且$\lambda_3 = aD - R_4$＜0 时，$E_4(0,1,1)$为 ESS。此时患者参与的风险损失小于收益，患者倾向于选择参与策略。由于患者选择参与，数据平台进行强监管时的信任收益变小，而成本大于强监管时的信任收益与弱监管时的信任损失之和，数据平台倾向于选择弱监管策略。同时，使用方不透明使用数据时的额外收益小于透明时的信任收益与不透明时的信任损失之和，使用方会选择透明使用策略。

结论 5：对于$E_6(1,0,1)$，当满足$\lambda_1 = G - T_4 - bF_2 - bL_1 - kF_1 + bkF_2$＜0 且$\lambda_2 = bF_2 - R_3 + bL_2 + kF_1 + bR_3 + bT_5 + kR_3 + kT_5 - bkF_2 - bkL_2 - bkR_3 - bkT_5 < 0$ 且$\lambda_3 = bD - T_6 - R_4$＜0 时，$E_6(1,0,1)$为 ESS。此时数据平台进行强监管时的成本小于其强监管时的信任收益、弱监管时的信任损失、强监管时的罚金所得之和，数据平台倾向于选择强监管策略。由于数据平台的强监管，患者的信任收益增大，患者的损失小于直接收益与信任收益之和，患者倾向于参与。对于使用方来说，由于不透明使用时的巨大额外收益的诱惑与逃避监管的投机心理，使用方倾向于选择不透明。这种情况对应于实际情境中，使用方不透明时潜在额外收益较大的情形。

结论 6：对于$E_8(1,1,1)$，当满足$\lambda_1 = G - T_4 - aL_1$＜0 且$\lambda_2 = R_3 - bF_2 - bL_2 - kF_1 - bR_3 - bT_5 - kR_3 - kT_5 + bkF_2 + bkL_2 + bkR_3 + bkT_5 < 0$ 且 $\lambda_3 = aD - T_6 - R_4 < 0$ 时，$E_8(1,1,1)$为 ESS。此时数据平台强监管的成本小于其强监管时的信任收益与弱监管时的信任损失之和，数据平台倾向于选择强监管策略。由于数据平台的强监管，患者的信任收益增大，患者的损失小于直接收益与信任收益之和，患者倾向于继续参与。对于使用方来说，采用透明使用策略获得的信任收益与采用不透明使用策略的信任损失之和大于其采用不透明使用策略时的额外收益与罚金之差，使用方倾向于选择透明使用策略。这种情况对应于实际情境中监管效果最好的情形。

4 仿真分析/Simulation and analysis

考虑未来我国健康大数据开放利用的实际情形，依据以下事实来模拟构建初始状态数据：

第一，我国非常重视健康大数据的开放利用，因此假设数据平台的信任收益T_4远大于G；

第二，随着大数据集的积累，数据使用的乘数效应与倍增效应会越发明显，假设各方的收益R_1~R_4都会不断增大，则使用方不透明使用的额外收益R_3也会增大的较快，即使用方面临不透明使用数据的利益诱惑会越来越大。

结合相关专家建议，设定各影响因素的基准初始值如表 4 所示，取$x = y = z = 0.5$，在 MATLAB 中进行仿真模拟，演化结果如图 3 a)所示，即系统的初始均衡状态为（强监管，

透明使用，参与）。依次研究以下两种预期状态变化：①R_3在不断增大；②R_3增大到一定程度后，各主要参数变化对消除不稳定波动的作用。

表 4 仿真模拟初始取值

Table 4 Initial values of the simulation

仿真初	R_1	G	T_4	k	F_1	F_2	L_1	C	R_2	R_3	T_5	L_2	L_3	R_4	D	a	b	T_6
始值	50	70	200	0.8	10	100	5	10	20	50	10	80	60	5	100	0.05	0.2	10

在上述设定的初始基准值的基础上，改变R_3的取值，依次取R_3=100,200,400，得到演化结果如图 3 b)所示。随着R_3的增大，数据平台的策略选择不变，仍然快速选择强监管策略。当R_3增大到某阈值（位于 100~200 之间）后，数据使用方与患者的策略选择出现波动，并且R_3越大，波动幅度越大。从演化过程中可以看出，当R_3突破该阈值后，数据使用方选择不透明使用策略的概率增大，进而患者隐私风险增大，逐步倾向于选择不参与，随着患者参与度的降低，使用方又逐步恢复透明，进而患者又恢复参与，由此往复循环，使用方与患者的策略选择演化结果形成如图 3 b)所示两条不稳定的波浪形线。

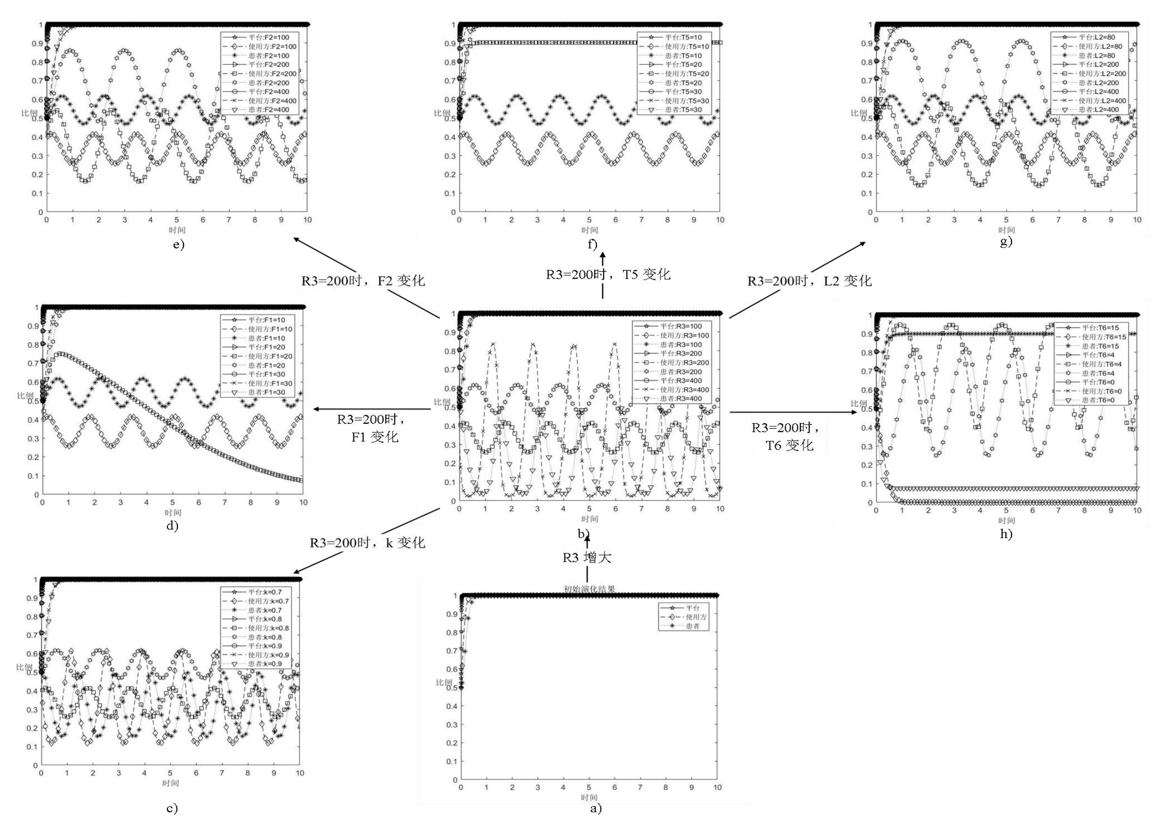


图 3 模型应用仿真过程与结果

Figure 3 Simulation process and results of model application

上述分析表明，当使用方不透明使用数据可获得的利益增大到一定程度（例如R_3=200）时，使用方和患者的策略选择会出现较大的波动，进一步研究其它各参数值变化对消除这种不稳定波动的作用。令表 4 中的R_3=200，其它参数取值保持不变，形成一组新的初始值。在此基础上，每次仅改变一个参数的值进行仿真分析。

图 3 c)为$k = 0.7$，$k = 0.8$，$k = 0.9$ 时的演化结果。k代表着数据平台的监管强度。相比较于$k = 0.7$ 时的演化曲线，当$k = 0.8$ 时使用方与患者的演化曲线上下波动幅度更小。当$k = 0.9$ 时，使用方快速选择透明使用策略，随之患者选择参与策略，系统重新恢复到（强监管，透明使用，参与）的均衡状态。由此可见，当k增大到一定阈值时（位于 0.8~0.9 之间），可以消除系统的不稳定波动，监管强度对数据使用的透明度和患者的参与度均具有正

向影响。同理，图 3 d)、e)分别反映出强监管时查到使用方不透明使用时的罚金额度、强监管时出现数据泄露事件后对使用方的罚金额度都对数据使用的透明度与患者参与具有正向影响。

图 3 f)为$T_5 = 10$，$T_5 = 20$，$T_5 = 30$ 时的演化结果。T_5代表着使用方透明使用的信任收益，也可视为对未来利益的考虑。当$T_5 = 20$ 时，使用方透明比例稳定在 0.9，当$T_5 = 30$ 时，使用方快速选择透明使用策略，随之患者选择参与策略，系统重新恢复到（强监管，透明使用，参与）的均衡状态。由此可见，使用方对透明使用时的信任收益对数据使用的透明度与患者的参与度具有正向影响作用。同理，图 3g)反映出使用方对不透明使用时的信任损失对数据使用的透明度与患者的参与度具有正向影响。

图 3 h)为$T_6 = 15$，$T_6 = 4$，$T_6 = 0$ 时的演化结果。T_6是数据平台强监管时患者参与得到的信任收益，代表着患者对数据平台的信任程度。当T_6过小（T_6=0）时，患者倾向于选择不参与，而使用方会倾向于选择透明使用；当T_6过大（$T_6 = 15$）时，由于患者出于对数据平台的信任倾向于选择参与，但此时反而助长了使用方的投机行为，使用方更倾向于选择不透明使用。由此可见，为促进健康大数据的长期可持续开放利用，在提升患者信任水平的同时，加强对使用方行为的制约更为重要。

5 结论与建议/Conclusions and suggestions

5.1 主要结论

开放利用实践表明，相关主体之间的信任已成为决定健康大数据开放利用成效的关键因素。聚焦数据平台、使用方、患者在考虑信任损益因素情况下的决策行为，基于开放利用过程分析，建立三方演化博弈模型，并通过 MATLAB 进行了仿真分析，以揭示信任对各方主体行为的影响机制。研究发现：

（1）获取患者信任是使用方透明使用数据的根本动力。使用方对透明使用数据时获得信任收益的考虑使系统避免了（监管/不监管，不透明，不参与）的不良局部均衡状态。使用方若失去患者信任，会导致大量患者选择退出，使用方可用数据量将急剧减少，获取患者信任是使用方选择透明使用策略的必要条件与根本动力。

（2）提升研究环境的安全性可以从根本上解决信任问题。当开放利用环境安全性低、患者感知风险极大时，会选择不参与；当开放利用环境安全性高、患者感知风险极小时，患者会选择参与。这两种极端情形下的患者的参与决策与数据平台和使用方的决策无关，与双方对信任损益的考虑程度也无关。

（3）加强监管与增大使用方的信任损益有助于实现良性的开放利用状态。当使用方不透明使用可获得的短期利益足够大时，会忽视信任损益而选择不透明使用策略，由此导致各方策略选择的不稳定波动。数据平台加强监管、提高罚金额度、使用方的信任损益提升到一定阈值后均可消除这种波动，促使系统恢复到（强监管，透明使用，参与）的均衡状态。

（4）提升患者对数据平台信任水平的同时，加强对使用方行为的制约更为重要。患者对数据平台充分信任时，会积极共享数据，但却会助长使用方不透明使用数据的投机行为；患者对数据平台高度不信任时，倾向于不参与，而此时数据使用方反而倾向于透明使用数据以挽回患者的信任。对使用方群体行为的治理是增进患者信任与促进健康大数据开放利用的根本所在。

5.2 政策建议

5.2.1 由国家主导推动健康大数据开放利用平台建设

国家主导推动健康大数据开放利用平台建设并对使用方群体实施强监管，充分发挥中央与地方政府的天然权威性优势，将有助于规范使用方行为与获取患者群体与数据供给机构的信任。我国公民对政府一直具有非常高的信任度[60]，公众对政府的信任会直接影响政府发布的信息是否为公众所信服、制定的方案是否能够被有效执行[61-62]。目前，我国已建立国家

人口健康科学数据中心、国家基因组科学数据中心、国家生物信息中心等国家主导的健康大数据开放利用平台，但存在数据质量较差、数据共享不足、数据利用率低等的问题[63]，还需打造更加权威（如通过 CoreTrustSeal 认证）和开放利用程度更高的健康大数据开放利用平台。建议由国家卫健委牵头、各地卫健委具体推动，建立国家级的健康大数据研究联盟，联合我国的健康大数据保管机构与研究机构等，汇聚、改进健康大数据集。与此同时，数据局、工信、公安等央地部门与卫健委协同，加强对使用方行为的规范力度，加强对使用方数据使用范围的审查，严格追查数据泄露源头并加大不规范使用行为的惩罚力度，在提升公众信任的同时大力规范使用方的行为，以更好的实现健康大数据的聚、通、用。

5.2.2 加强可信研究环境建设并积极推广应用

通过技术手段增强健康大数据访问的安全性，是降低患者隐私泄露风险的“硬手段”，有利于获得患者群体与数据供给机构的信任。可信研究环境是基于网络技术与管理手段，为科研人员访问敏感数据而设计的安全环境。它消除了在不同用户之间物理共享数据的需要，通过安全的人、安全的项目、安全的设置、安全的数据、安全的输出等 5 大方面的安全技术与管理保障，真正实现“数据可用不可见”，进而保障科研人员在快捷访问、分析所需数据的同时，规避因数据下载而导致的数据泄露等安全责任与风险，保障数据安全，同时建立公众信任，实现信息治理、透明度和安全性的最高标准[64]。从项目实施角度看，TRE 在确保隐私保护等方面的数据安全的同时，最大限度地利用高性能计算并减少数据传输、存储的费用和安全风险，进而有效推动数据利用相关合作项目的达成。建议健康大数据开放利用平台加强可信研究环境建设，研发或引进相关技术，加强相关基础设施建设，结合数据分类分级体系将高度敏感的数据在 TRE 中进行分析研究，减免高敏感数据的流动，在保障数据尽可能安全的前提下平衡利用成本与效益。

5.2.3 推行数据使用登记制度并建立标准体系

数据使用登记制度的实施可以有效减少数据使用方的的不透明使用行为。明确数据的使用方与使用目的等信息，促进健康医疗大数据使用的主体透明、目的透明及过程透明，将数据使用方置于“明处”，接受公众审查与监督，可以促进使用方透明使用数据与提升公众信任。推动数据使用方发布数据使用登记册是实施数据使用登记制度的具体方式。数据使用登记册，也称为数据发布登记册或批准研究清单，是机构为研究、创新和服务评估目的向批准的个人和组织提供安全访问的数据的公共记录，它通常包含有关被访问数据的类型、用途、批准日期和使用（或接收）数据的组织名称等信息[65]。数据使用登记册是提高数据使用透明度为基石，有助于建立公众信心、增进公众信任，有助于向社会展示使用数据的价值和好处、便于科研人员和科研资助机构了解更多可用数据集的信息、便于相关主管部门了解数据的使用情况等，具有广泛的社会效益。建议我国参考英国健康数据中心的数据使用登记制度及其标准化工作实践[65]，结合我国实际推进我国数据使用登记工作有序开展。

5.2.4 有效引导公众参与健康大数据协同治理过程

健康医疗大数据本质上是个人健康数据的宏观集合，涉及最广泛的公众权益，让公众参与健康大数据的开放利用决策并行使监督权，有利于规范使用方的使用行为并增进患者与数据供给机构的信任。政府要引导社会各界重视公众价值，创造公众参与氛围，在健康大数据相关政策制定时就把公众参与协同治理的理念融入其中，进行自上而下的观念渗透；要授予公众参与权利，提供公众参与机会，通过组织机构设置与公共活动发布为公众提供合法合规参与渠道，建议各地方卫健委下设由公众组成的公共咨询委员会并公开透明的通过选举来确定成员，充分发挥公众的建议与监督职能；要完善制度规范体系，引导公众有序参与。例如制定协同参与的准入政策，确保参与公众是具有代表性的；制定具体参与过程的制度规范，明晰公众参与的目标以及工作内容并规范其参与行为；制定公众参与结果的反馈制度等。最后，提供资源支持，化解公众参与障碍，例如提供知识支持、岗前培训以及物质激励等。

5.3 研究局限与展望

在对英国健康医疗大数据开放利用实践进行分析的基础上，重点考虑了使用方的透明度建设与各方的信任损益，建立了三方演化模型，对健康大数据开放利用主体之间决策行为的相互影响进行了有益探索。但本研究为便于与其它收益或成本进行计算，将信任损益视为各方对未来利益的考虑，仅对信任损益进行了简单的字母设定与量化赋值，探讨了单个变量的变化对演化结果的影响，存在一定的局限性。未来可对信任损益量化的方法、仿真取值的设定进行更科学的探讨，对多个变量同时变化对演化结果的影响进一步讨论。另外，文章仅围绕英国实践及其信任机制进行研究，后续可以对不同国家在健康大数据开放利用中的信任机制差异进行更加深入的研究和对比分析，以资为我国在全球健康数据治理中的参与提供更多的参考。

参考文献：

[1] 国家卫生健康委员会. 关于印发国家健康医疗大数据标准、安全和服务管理办法（试行）的通知[EB/OL]. [2024-07-22]. http://www.nhc.gov.cn/mohwsbwstjxxzx/s8553/201809/f346909ef17e41499ab766890a34bff7.shtml.(National Health Commission of the People’s Republic of China. Notice on printing and distributing the “National healthcare big data standards, security, and service management measures (trial)”[EB/OL]. [2024-07-22]. http://www.nhc.gov.cn/mohwsbwstjxxzx/s8553/201809/f346909ef17e41499ab766890a34bff7.shtml.)

[2] HDR UK. Our strategy[EB/OL]. [2024-07-25]. https://www.hdruk.ac.uk/about-us/our-strategy/.

[3] 高富平. 论医疗数据权利配置──医疗数据开放利用法律框架[J]. 现代法学, 2020, 42(04): 52-68.(GAO F P. An entitlement of clinical data: the lega framework of clinical data sharing[J]. Modern law science, 2020, 42(04): 52-68.)

[4] 李雅琴. 数字治理视域下健康数据利用权益制度构建[J]. 湖北大学学报(哲学社会科学版), 2024, 51(02): 163-173.(LI Y Q. Construction of the rights and interests system for the use of health data from the perspective of digital governance[J]. Journal of Hubei University(philosophy and social science), 2024, 51(02): 163-173.)

[5] 任颖. 医疗数据使用权的私法保障与边界厘定[J]. 政法论丛, 2023(06): 99-110.(REN Y. Private law protection of the right to use medical data[J]. Journal of political science and law, 2023(06): 99-110.)

[6] 汤建华. 健康医疗大数据共享风险及法治因应[J]. 学习与实践, 2023(08): 54-60.(TANG J H. Risks and legal responses to healthcare big data sharing[J]. Study and practice, 2023(08): 54-60.)

[7] 李嘉兴, 王雷, 宋士杰等. 重大突发公共卫生事件驱动的医疗数据开放治理模式研究[J]. 图书情报工作, 2022, 66(04): 23-32.(LI J X, WANG L, SONG S J, et al. Research on open governance model for healthcare data driven by major public health emergencies[J]. Library and information service, 2022, 66(04): 23-32.)

[8] 翟运开, 郭瑞芳, 王宇等. 数据生命周期视角下的医疗健康大数据质量评价研究[J]. 现代情报, 2024, 44(01): 116-129.(ZHAI Y K, GUO R F, WANG Y, et al. Research on quality evaluation of medical and health big data from the perspective of data life cycle[J]. Journal of modern information, 2024, 44(01): 116-129.)

[9] 郭宇, 张传洋, 刘梦婷等. 用户信息焦虑纾解驱动医疗健康大数据价值实现影响因素与路径研究[J]. 图书情报工作, 2023, 67(15): 25-34.(GUO Y, ZHANG C Y, LIU M T, et al. Research on influencing factors and paths for the reduction of users’ information anxiety drives the realization of the value of medical and health big data[J]. Library and information service, 2023, 67(15): 25-34.)

[10] 翟运开, 宋欣, 王宇. 医疗健康大数据资产价值实现路径分析——基于信息生态系统理

论[J]. 技术经济, 2023, 42(11): 178-190.(ZHAI Y K, SONG X, WANG Y. Analysis of the value realization path of medical and health big data assets: based on information ecosystem theory[J]. Journal of technology economics, 2023, 42(11): 178-190.)
[11] SINGH K, LI S, JAHNKE I, et al. Improving big data governance in healthcare institutions: user experience research for honest broker based application to access healthcare big data[J]. Behaviour & information technology, 2023, 43(6): 1067-1095.
[12] JIANG R, CHEN X, YU Y, et al. Risk and UCON-based access control model for healthcare big data[J]. Journal of big data, 2023, 10(01): 1-28.
[13] CHEN H, SONG Z, YANG F. Storage method for medical and health big data based on distributed sensor network[J]. Journal of sensors, 2023, 2(3): 1-10.
[14] SAROSH P, PARAH S A, MALIK B A, et al. Real-time medical data security solution for smart healthcare[J]. IEEE transactions on industrial informatics, 2022, 19(07): 8137-8147.
[15] ZHU L, ZHOU Y, FAN Q, et al. Enabling efficient and secure health data sharing for healthcare IoT systems[J]. Future generation computer systems, 2023, 149: 304-316.
[16] MAHAJAN B, JUNNARKAR A. Smart healthcare system using integrated and lightweight ECC with private blockchain for multimedia medical data processing[J]. Multimedia tools and applications, 2023, 82(28): 44335-44358.
[17] ALMALAWI A, KHAN A I, ALSOLAMI F, et al. Managing security of healthcare data for a modern healthcare system[J]. Sensors, 2023, 23(07): 3612.
[18] SARANYA K, PREMALATHA K. Multi attribute case based privacy-preserving for healthcare transactional data using cryptography[J]. Intelligent automation & soft computing, 2023, 35(02): 2029-2042.
[19] KERASIDOU X, MALONE M, DALY A, et al. Machine learning models, trusted research environments and UK health data: ensuring a safe and beneficial future for AI development in healthcare[J]. Journal of medical ethics, 2023, 49: 838-843.
[20] BRAUNACK-MAYER A, CAROLAN L, STREET J, et al. Ethical issues in big data: a qualitative study comparing responses in the health and higher education sectors[J]. Plos one, 2023, 18(04): e0282285.
[21] RASHI T, YOM-TOV E. Ethics of medical archival internet research data[J]. Journal of medical internet research, 2023, 25: e43754.
[22] GIUFFRÈ M, SHUNG D L. Harnessing the power of synthetic data in healthcare: innovation, application, and privacy[J]. NPJ digital medicine, 2023, 6(01): 186.
[23] RENNIE S, ATUIRE C, MTANDE T, et al. Public health research using cell phone derived mobility data in Sub-Saharan Africa: ethical issues[J]. South African journal of science, 2023, 119(5-6): 1-7.
[24] TOMMEL J, KENIS D, LAMBRECHTS N, et al. Personal genomes in practice: exploring citizen and healthcare professionals' perspectives on personalized genomic medicine and personal health data spaces using a mixed-methods design[J]. Genes, 2023, 14(04): 786.
[25] SCHICKHARDT C, WINKLER E, SAX U, et al. Data infrastructures for health research: ethical framework and legal implementation[J]. Bundesgesundheitsblatt, gesundheitsforschung, gesundheitsschutz, 2023, 66(02): 160-167.
[26] VEZYRIDIS P, TIMMONS S. Understanding the care.data conundrum: new information flows for economic growth[J]. Big data & society, 2017, 4(01): 1-12.
[27] PRESSER L, HRUSKOVA M, ROWBOTTOM H, et al. Care.data and access to UK health records: patient privacy and public trust[J]. Technology science, 2015, 2015081103: 1-35.
[28] CARTER P, LAURIE G T, DIXON-WOODS M. The social licence for research: why care.data ran into trouble[J]. Journal of medical ethics, 2015, 41(05): 404-409.
[29] 苏强, 季荔. 基于随机演化博弈的医疗数据共享协调机制研究[J]. 情报科学, 2023, 41(09): 37-47.(SU Q, JI L. The coordination mechanism of medical data sharing based on stochastic evolutionary game[J]. Information science, 2023, 41(09): 37-47.)
[30] 韩普, 顾亮, 张嘉明. 隐私保护视角下医疗数据共享意愿研究——基于三方演化博弈分析[J]. 现代情报, 2021, 41(03): 148-158.(HAN P, GU L,ZHANG J M. Research on willingness to share medical data from perspective of privacy protection——based on tripartite evolutionary game analysis[J]. Journal of modern information, 2021, 41(03): 148-158.)
[31] 韩普, 张嘉明. 考虑隐私保护的医疗数据共享意愿研究——基于演化博弈的视角[J]. 现

代情报, 2021, 41(01): 111-120.(HAN P, ZHANG J M. Research on the willingness of medical data sharing considering privacy protection——from the perspective of evolutionary games[J]. Journal of modern information, 2021, 41(01): 111-120.)
[32] BAIER A. Trust and antitrust[J]. Ethics, 1986, 96(02): 231-60.
[33] HAWLEY K. Trust, distrust and commitment[J]. Noûs, 2014, 48(01): 1-20.
[34] O'NEILL O. Trust, trustworthiness, and accountability[J].Oxford University press, 2014: 172-190.
[35] LHOUSNI S, DAOUDI F, BELMOKHTAR I, et al. Patients' knowledge and attitude toward biobanks in Eastern Morocco[J]. Biopreservation and biobanking, 2020, 18(03): 189-195.
[36] GASKELL G, GOTTWEIS H. Biobanks need publicity[J]. Nature, 2011, 471(7337): 159-160.
[37] GASKELL G, ALLANSDOTTIR A, ALLUM N, et al. The 2010 Eurobarometer on the life sciences[J]. Nature biotechnology, 2011, 29(02): 113-114.
[38] PAWLIKOWSKI J, WIECHETEK M, MAJCHROWSKA A. Associations between the willingness to donate samples to biobanks and selected psychological variables[J]. International journal of environmental research and public health, 2022, 19(05): 2552.
[39] SANDERSON S C, BROTHERS K B, MERCALDO N D, et al. Public attitudes toward consent and data sharing in biobank research: a large multi-site experimental survey in the US[J]. The American journal of human genetics, 2017, 100(03): 414-427.
[40] MILLS E J, SEELY D, RACHLIS B, et al. Barriers to participation in clinical trials of cancer: a meta-analysis and systematic review of patient-reported factors[J]. The lancet oncology, 2006, 7(02): 141-148.
[41] MIDDLETON A, MILNE R, ALMARRI M A, et al. Global public perceptions of genomic data sharing: what shapes the willingness to donate DNA and health data?[J]. The American journal of human genetics, 2020, 107(04): 743-752.
[42] DIVE L, CRITCHLEY C, OTLOWSKI M, et al. Public trust and global biobank networks[J]. BMC medical ethics, 2020, 21(01): 1-9.
[43] MILNE R, MORLEY K I, ALMARRI M A, et al. Demonstrating trustworthiness when collecting and sharing genomic data: public views across 22 countries[J]. Genome medicine, 2021, 13(01): 92.
[44] AITKEN M, JORRE J, PAGLIARI C, et al. Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies[J]. BMC medical ethics, 2016, 17(01): 1-24.
[45] LEMKE A A, WOLF W A, HEBERT-BEIRNE J, et al. Public and biobank participant attitudes toward genetic research participation and data sharing[J]. Public health genomics, 2010, 13(06): 368-377.
[46] SHABANI M, BEZUIDENHOUT L, BORRY P. Attitudes of research participants and the general public towards genomic data sharing: a systematic literature review[J]. Expert review of molecular diagnostics, 2014, 14(08): 1053-1065.
[47] BALLANTYNE A. How should we think about clinical data ownership?[J]. Journal of medical ethics, 2020, 46(05): 289-294.
[48] 陈怡. 健康医疗数据共享与个人信息保护研究[J]. 情报杂志, 2023, 42(05): 192-199. (CHEN Y. Research on healthcare data sharing and personal information protection[J]. Journal of intelligence, 2023, 42(05): 192-199.)
[49] VOIGT P, BUSSCHE A. The EU general data protection regulation (GDPR), a practical guide[M]. Cham: Springer international publishing, 2017, 10(3152676): 10-5555.
[50] 高富平, 李群涛. "个人信息权益"的民法新定位[J]. 东岳论丛, 2023, 44(06): 166-176.(GAO F P, LI Q T. The new civil law positioning of "personal information rights and interests"[J]. Dongyue tribune, 2023, 44(06): 166-176.)
[51] WEITZMAN E R, KELEMEN S, KACI L, et al. Willingness to share personal health record data for care improvement and public health: a survey of experienced personal health record users[J]. BMC medical informatics and decision making, 2012, 12(01): 1-10.
[52] MEDFORD-DAVIS N, CHANG L, RHODES V. Health information exchange: what do patients want?[J]. Health informatics journal, 2017, 23(04): 268-278.
[53] ATHEY S, CATALINI C, TUCKER C. The digital privacy paradox: small money, small costs, small talk[R]. National Bureau of economic research, 2017.

[54] MÉJEAN C, EDELENYI F, TOUVIER M, et al. Motives for participating in a web-based nutrition cohort according to sociodemographic, lifestyle, and health characteristics: the NutriNet-Sante cohort study[J]. Journal of medical internet research, 2014, 16(08): e189.
[55] HOWE N, GILES E, NEWBURY-BIRCH D, et al. Systematic review of participants' attitudes towards data sharing: a thematic synthesis[J]. Journal of health services research & policy, 2018, 23(02): 123-133.
[56] 邱均平, 徐中阳, 陈锡慧. 基于三方演化博弈的在线健康社区用户隐私披露行为研究[J]. 情报理论与实践, 2023, 46(01): 24-36.(QIU J P, XU Z Y, CHEN X H. Research on user privacy disclosure behavior in online health communities based on tripartite evolutionary game[J]. Information studies: theory & application, 2023, 46(01): 24-36.)
[57] RITZBERGER K, WEIBULL J W. Evolutionary selection in normal-form games[J]. Econometrica, 1995, 63(6): 1371-1399.
[58] FRIEDMAN D. Evolutionary games in economics[J]. Econometrica, 1991, 59(03):637-666.
[59] 朱立龙, 荣俊美, 张思意. 政府奖惩机制下药品安全质量强监管三方演化博弈及仿真分析[J]. 中国管理科学, 2021, 29(11): 55-67.(ZHU L L, RONG J M, ZHANG S Y. Three-party evolutionary game and simulation analysis of drug quality supervision under the government rewardand and punishment mechanism[J]. Chinese journal of management science, 2021, 29(11): 55-67.)
[60] 爱德曼国际公关公司与清华大学国家形象传播研究中心. 2023 年爱德曼信任度调查中国报告[EB/OL]. [2024-08-06]. https://file.digitaling.com/eImg/uimages/20230403/1680510992119917.pdf. (Edelman Public Relations Worldwide, the National Image Communication and Research Center at Tsinghua University. 2023 Edelman trust barometer China report[EB/OL]. [2024-08-06]. https://file.digitaling.com/eImg/uimages/20230403/1680510992119917.pdf.)
[61] FANCOURT D, STEPTOE A, WRIGHT L. The cummings effect: politics, trust, and behaviours during the COVID-19 pandemic[J]. The lancet, 2020, 396(10249): 464-465.
[62] KREPS S E, KRINER D L. Model uncertainty, political contestation, and public trust in science: evidence from the COVID-19 pandemic[J]. Science advances, 2020, 6(43): eabd4563.
[63] 吴思竹, 王安然, 修晓蕾, 等. 欧美生物医学科学数据中心建设及启示[J]. 数字图书馆论坛, 2022(04): 2-10.(WU S Z, WANG A R, XIU X L, et al. Construction and implications of biomedical scientific data centers in Europe and the United States[J]. Digital library forum, 2022(04): 2-10.)
[64] RUTH M. A pioneering programme to establish a national infrastructure for health data science[EB/OL]. [2024-08-02]. https://www.hdruk.ac.uk/news/a-pioneering-programme-to-establish-a-national-infrastructure-for-health-data-science/.
[65] KARRAR N, KHAN S K, MANOHAR S, et al. Improving transparency in the use of health data for research: recommendations for a data use register standard[EB/OL]. [2024-08-08]. https://zenodo.org/record/5902743#.Y19-A2DjzuE.

作者贡献说明/Author contributions:
王丹丹：确定论文选题与整体研究框架，撰写论文，修改论文；
解世程：撰写论文，修改论文；
程英英：修改论文。

Evolutionary Game Analysis of Open Utilization of Healthcare Big Data Considering Trust Gains and Losses

Wang Dandan, Xie Shicheng, Cheng Yingying
Business School of Henan University of Science and Technology, Luoyang 471023

Abstract: [Purpose/Significance] Trust among relevant stakeholders has become a key factor in determining the successful open utilization of healthcare big data. This study investigates the impact mechanism of trust on the stakeholder behavior, considering the factors of trust gains and losses, providing forward-looking references for improving the effectiveness of healthcare big data open use in China. **[Methods/Process]** Focusing on the decision-making behavior of the data platform, users, and patients, and considering factors of trust gains and losses, this study established a tripartite evolutionary game model based on the analysis of the open utilization process. Simulation analysis was conducted using MATLAB, and countermeasure suggestions were proposed based on the equilibrium point analysis and simulation results. **[Results/Conclusion]** Gaining the trust of patients is fundamental for users to transparently use data. Enhancing the security of the research environment can fundamentally solve trust issues. Strengthening regulation and increasing the trust gains and losses for users contribute to a positive state of open utilization. While raising the trust level of patients in data platforms, it is more crucial to intensify constraints on user behavior. China may consider building a nationally-led platform for the open utilization of healthcare big data, strengthening the construction of a trusted research environment and actively promoting its application, implementing a data use registration system, and establishing a standard system, for effectively guiding the public to participate the collaborative governance of healthcare big data.

Keywords: Healthcare big data Trust Transparency Open Utilization Evolutionary Game